FEBS 29157 FEBS Letters 579 (2005) 877–884

Minireview

Origin recognition and the chromosome cycle

Bruce Stillman

Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA

Received 6 December 2004; revised 10 December 2004; accepted 10 December 2004

Available online 18 December 2004

Edited by Gunnar von Heijne and Anders Liljas

Abstract Prior to the initiation of DNA replication, chromosomes must establish a biochemical mark that permits the recruitment in S phase of the DNA replication machinery that copies DNA. The process of chromosome replication in eukaryotes also must be coordinated with segregation of the duplicated chromosomes to daughter cells during mitosis. Protein complexes that utilize ATP coordinate events at origins of DNA replication and later they participate in the initiation of DNA replication. In eukaryotes, some of these proteins also play a part in later processes that ensure accurate inheritance of chromosomes in mitosis, including spindle attachment of chromosomes, accurate duplication of centrosomes and cytokinesis. A perspective of how ATP-dependent proteins accomplish this task in eukaryotes is discussed.

Keywords: DNA replication; Origin recognition complex; AAA^+ ATPases

1. Introduction

All chromosomes in eukaryotic cells must be duplicated and then segregated to daughter cells to preserve the genetic heritage of the species and to produce the vast majority of cells that make up an individual organism. In eukaryotes, and particularly in multi-cellular eukaryotes, mistakes during mitotic chromosome inheritance can have devastating consequences to the individual because DNA damage can lead to the establishment of clones of cells that proliferate uncontrollably, such as occurs in cancer. To ensure accurate inheritance of genetic information, chromosomes must be precisely copied only once during the cell cycle and then evenly segregated to progeny cells. Coordination of the processes of chromosome duplication and segregation is therefore a key aspect of the cell division cycle. Overall, the two processes are regulated by the cyclin-dependent protein kinases that coordinate cell cycle progression. These and other protein kinases directly regulate complex machineries that participate in both chromosome replication and segregation. This perspective discusses some of the issues involved in the inheritance of DNA in eukaryotes and highlights similarities to inheritance of virus, bacterial and archaeal chromosomes.

2. Establishing the competence of chromosomes to replicate

Studies in the early 1970s suggested that in eukaryotic cells, chromosomes are somehow marked in the G1 phase and this mark renders them competent for initiation of DNA replication (for example, see [1]). G1 chromosomes, however, do not initiate DNA replication because the cells lack activators that trigger entry into S phase. These studies also demonstrated that once replicated, the competent state of chromosomes is lost and they have to pass through mitosis before they can regain this marked state. We now know that the competency of eukaryotic chromosomes to replicate is established by the assembly of a multi-protein complex (a pre-replicative complex; pre-RC) at individual origins of DNA replication [2,3]. The pre-RC consists of the origin recognition complex (ORC) that is the primary recognition protein for origins of DNA replication [4]. ORC binding allows recruitment of other known pre-RC proteins such as Cdc6 [5–9], Cdt1 [10,11] and the mini-chromosome maintenance proteins (Mcm2p, Mcm3p, Mcm4p, Mcm5p, Mcm6p and Mcm7p; [12–18]).

The establishment of the competent state occurs in a period of the cell cycle when cyclin-dependent protein kinases are inactive. For example, in rapidly proliferating cells, the pre-RC is established during exit from mitosis and the marked chromosomes are inherited into the new daughter cells [19]. In quiescent cells that are stimulated to re-enter the cell cycle, pre-RC formation occurs in late G1 phase [18]. In either case, the competent state of chromosomes is activated and the DNA replication machinery acts in a temporally specific manner at each origin of DNA replication to replicate a segment of DNA (called a replicon). During this period, the pre-RC is disrupted and individual components of the pre-RC are either destroyed or dissociate from the DNA. Chromosomes therefore lose the ability to re-initiate DNA replication because the pre-RC that is necessary for initiation of DNA replication is not present [2,3]. The presence of cyclin-dependent protein kinases prevents new pre-RC complex formation. After chromosome segregation has been completed, the whole process may begin again.

While the biochemistry of the inheritance of chromosomes is still an active area of study, much of what we know about DNA replication in eukaryotic cells derives from studies on double-stranded DNA viruses that infect mammalian cells. Thus, it is informative to understand how origin recognition and initiation occurs in viral systems.

E-mail address: stillman@cshl.edu (B. Stillman).

0014-5793/$30.00
doi:10.1016/j.febslet.2004.12.011

3. Origin recognition: SV40 and related viruses

Viruses with double stranded DNA genomes, such as Simian Virus 40 (SV40), infect primate cells and produce many progeny DNA molecules in a single infection. Thus, they ignore the once per cell division rule imposed on cell chromosomes, even though the viruses utilize many of the cellular proteins that are required for duplication of the cell chromosome. In the case of SV40, the best-studied system, SV40 T antigen (T Ag) is the only virus-encoded protein required for viral DNA replication and it plays essential roles in all stages of DNA replication [20].

T Ag binds to four repeated DNA sequence motifs within the genetically defined origin of DNA replication [21]. In contrast to assembly of the pre-RC at cellular origins of DNA replication that is inhibited by active cyclin-dependent protein kinases, binding of T Ag to the origin of DNA replication requires cyclin-dependent protein kinase activity [22]. This requirement may enable T Ag to bind only when cells are proliferating and have committed to produce the cellular replication proteins on which virus DNA replication relies.

In the presence of ATP, the T Ag multimerizes and eventually forms two hexamers that remain associated with each other at the replication fork (Fig. 1) [23]. Each hexamer binds to the DNA in a non-sequence-specific manner and has DNA helicase activity. Thus, the duplex DNA is threaded through the structure, rather than the helicase moving along the DNA.

A related protein encoded by papillomaviruses, called the E1 protein, binds to specific sequences at the virus origin of DNA replication and forms double-hexamers in an ATP dependent process (Fig. 1) [24–26]. In the case of the papillomavirus E1 protein, DNA sequence specific binding of the protein to the virus origin of DNA replication is aided by a second virus encoded protein called E2 [27,28]. The E2 protein binds with high affinity to specific recognition sequences in the viral DNA and a separate domain of the E2 protein interacts with E1 monomers and therefore both bind cooperatively to the origin. The presence of E2 protein prevents E1 protein from forming

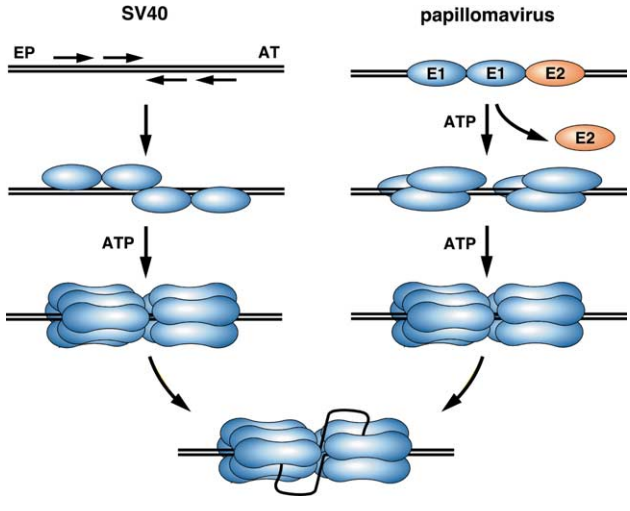

Fig. 1. Virus Origin recognition proteins. The SV40 T antigen (left) and the papillomavirus E1 and E2 proteins each recognize their respective origins as monomers in a DNA sequence specific manner and then form two hexamer helicase structures at the origin in an ATP-dependent manner. SV40 T Ag structure is modeled from [32]. The E1 and E2 mechanism is described in more detail in [24].

hexamer structures, however during the transition from the monomer to the hexamer form of E1, interaction between the E2 protein is disrupted by E1 binding to ATP, forcing the release of E2 from E1 and the origin of replication [26,28,29]. Thus, the E2 protein acts both as a positive chaperone for E1 binding to the origin, but also as a negative regulator of E1 helicase activity. In this way, E1 binds to the origin of DNA replication and assembles the helicase activity at the origin rather than randomly along the chromosome. An ATP-driven event causes the transition from the origin recognition state to the helicase state. Recent crystal structures of an E1–E2 complex suggest that ATP causes an allosteric change in the structure of E1 that is incompatible with E2 interaction, but compatible with hexamer formation [29]. Once formed, the E1 AAA^+ helicase is believed to assume a structure similar to the SV40 T Ag helicase.

Recent electro microscopy and X-ray crystallography studies on the structure of SV40 T Ag have suggested mechanisms for origin recognition, ATP-driven multimerization and DNA helicase activity [30–32]. T Ag has a small origin recognition domain in the amino terminal region of the protein that binds to the repeated sequences within the virus origin of DNA replication [33]. Four such recognition sites exist in the origin and based on what is known about the steps involved to form the papillomavirus E1 hexamer, it is likely that two of the T Ag molecules are precursors for the assembly of a single hexamer. Since four molecules of T Ag initially recognize the origin of DNA, two hexamers will form in the presence of ATP. These two hexamers most likely remain bound together [23]. Recent determination of the structure of the helicase domain shows a central channel in the helicase that can accommodate DNA and a side channel that could accommodate the separated single stranded DNA [31,32]. The joined helicase domains point away from each other and therefore away from the center of the symmetrical origin. This orientation is consistent with the two hexamers pumping the DNA though the structure, unwinding the DNA and allowing DNA synthesis primed by the polymerase α-primase complex that binds to T Ag (Fig. 1).

The T Ag helicase structure revealed a surprising potential mode of ATP driven DNA unwinding. Previous structural and biochemical studies on many multi-subunit AAA^+ ATPases, including the DNA replication clamp loader RFC, suggested that ATP hydrolysis on one subunit is followed by hydrolysis on an adjacent subunit [34,35]. In contrast, the T Ag structure in its different nucleotide bound forms suggests that ATP hydrolysis occurs coordinately on all T Ag subunits in the hexameric helicase. Furthermore, the ATP bound form shows that a β-hairpin structure in the central channel of the molecule is in a closed conformation, whereas the ADP bound form shows the central channel and the β-hairpin open [32]. This has led to a model in which the β-hairpin interacts with the DNA and in the presence of ATP moves the DNA through the central channel, unwinding it and pumping the single stranded DNA out through the side channels (Fig. 1).

A remarkably similar structure exists in the hexameric MCM proteins. Eukaryotes have six essential MCM proteins that form a hexameric structure [36], whereas many archaeal species have a single MCM protein that forms a double-hexamer made of twelve identical subunits [37–39]. X-ray crystallography of the amino-terminal half of the protein shows it to be a barrel shaped double hexamer with a central channel that could accommodate double stranded DNA. Furthermore,

the MCM fragment contained a β-hairpin in the channel similar to that found in SV40 T Ag. Thus, it is expected that the MCM helicase might function like SV40 T Ag helicase by pumping the DNA through its central channel.

Unlike T Ag, where two hexamers sit at the two, joined DNA replication forks, the eukaryotic MCM proteins are widely distributed along the chromosomes before DNA replication and there is little evidence that the bulk of MCM protein is associated with active DNA synthesis. A model to explain the multiple MCM proteins that are loaded along the unreplicated portion of the DNA suggests that the MCM proteins may pump DNA toward the replication fork, as well as participate in unwinding the DNA at the fork itself [40]. Thus, the MCM proteins would function to translocate DNA as well as act as a DNA helicase. Hexameric tranlocases that utilize ATP have been observed to promote branch migration of Holliday junctions at sites of DNA recombination.

Another possible role of the MCM proteins that are bound to chromatin away from the DNA replication fork itself may be to remove histones from the DNA before the DNA enters into the DNA synthesis machinery. For example, if the two dimers of histones H2A.H2B that are present on each nucleosome were removed from the H3.H4 tetramer histone core, one negative supercoil in the nucleosomal DNA would be relaxed. In topologically constrained DNA, such as occurs in eukaryotic chromosomes, a change in twist in the nucleosomal DNA by the loss of H2A.H2B dimers from one nucleosome could promote compensatory unwinding of the DNA by ten base pairs. If multiple nucleosomes were disrupted as the DNA translocated, a substantial region of the DNA may already be unwound, or at least be energetically favorable for unwinding, prior to entry of the DNA synthesis machinery.

4. Origin recognition in eukaryotes

Unlike the small DNA viruses that have a single origin of DNA replication, eukaryotic cells have multiple large chromosomes that require many start sites for replication so that they can complete the task of duplicating the genome within a single S phase of the cell cycle. It was not obvious that the start sites for initiation would be DNA sequence specific and indeed many studies had demonstrated that in *Xenopus* oocytes, any DNA, even bacterial phage and plasmid DNAs, would replicate only once per cell division cycle [41]. The concept emerged that in vertebrates the start sites for initiation of DNA replication were not genetically determined, a situation different to what was known to exist in bacterial and plasmid genomes.

This concept contrasted with DNA transformation experiments in *S. cerevisiae* where fortuitously it was discovered that some DNA sequences caused high frequency transformation of selected gene markers [42]. Later studies demonstrated that these autonomously replicating sequences (ARSs) correspond to start sites for initiation of DNA replication [43]. Genetic analysis of the sequences revealed a multi-domain structure of the genetically determined sequences, called a replicator [44]. The origin sequences in the *S. cerevisiae* genome consist of an essential A element and multiple important B elements that collectively are required for origin function (Fig. 2). In other yeast species, such as *S. pombe*, genetically defined ARS sequences exist, but detailed analyses of the DNA sequences have not revealed small DNA elements like those that exist in *S. cerevisiae*. Rather, ORC binds in an ATP independent manner to multiple, but redundant A.T-rich sequence blocks [45–48]. Interestingly, *S. pombe* ORC has a subunit (Orc4p) that has an A.T hook domain associated with the AAA^+ domain of the Orc4 subunit [49]. It is possible that *S. pombe* ORC represents an anomaly, but it is also possible that origin determination in higher eukaryotes involves ORC interacting with separate A.T hook domain proteins.

The identification of specific sequences at origins of DNA replication in *S. cerevisiae* resulted in the identification of origin binding proteins. The first such protein that was shown to be physiologically relevant was the Abf1p that bound to the B3 element in the *ARS1* origin [50]. This protein functions as an enhancer of DNA replication, analogous to DNA-sequence specific enhancers of gene transcription and indeed in other contexts, Abf1p can either activate or repress transcription. A protein complex containing the Myb protein may perform a similar role in *Drosophila* [51]. But the critical protein that binds to all known origins of DNA replication in all eukaryotes is the ORC [4]. The binding site for ORC within yeast origins of DNA replication corresponds to the region where origins are hypersensitive to nuclease digestion in chromatin during S and G2 phases of the cell cycle and are therefore relatively protected during the G1 phase of the cell cycle [52,53]. This observation led to the idea that a pre-replicative complex (pre-RC) is bound to origins of replication in the G1 phase of the cell cycle [54]. The pre-RC is almost certainly the mark discussed above that renders chromosomes competent for DNA replication and is found on G1 phase chromosomes.

A six subunit protein complex, ORC is conserved in all eukaryotes and the largest subunit, Orc1p is even conserved in archaeal species (see below; [2,55]). All six subunits are essential in yeast, but only the Orcl-5 subunits are essential for DNA sequence-specific binding to origins in vitro [56]. ORC was also found to bind to DNA sequences in the yeast genome that normally do not function as origins of DNA replication, but act as silencers of gene transcription [57]. This raised the possibility that ORC may have other roles in the cell, a scenario that has been realized in animal cells (see below).

ORC binds to origins of DNA replication in an ATP dependent manner, a highly unusual circumstance for sequence-specific DNA binding proteins [4]. Some subunits of ORC have a AAA^+ domain that either bind ATP or bind and hydrolyze ATP [58–61]. Furthermore, the ATPase activity of ORC is regulated by DNA [58]. ATPase activity is inhibited when ORC is bound to origins of DNA replication and is activated by single stranded DNA, suggesting that ORC plays an active role in origin determination and function and does not just function as a landing pad for other replication proteins.

In yeast, all ORC subunits are bound to the chromatin throughout the cell division cycle, with some subunits being phosphorylated in a cell cycle dependent manner by S-phase cyclin-CDKs [62–65]. In contrast, the mammalian ORC is highly dynamic during the cell cycle, with Orc1 being phosphorylated and degraded by ubiquitin-mediated proteolysis during the G1 to S phase transition in human cells or monoubiquityled and released from chromatin in hamster cells [66–73]. Moreover, the Orc2 subunit is bound to chromatin in the G1 phase of the cell cycle in mammalian cells, but is gradually removed during S phase and ends up mostly on centromeric heterochromatin in mitosis [74]. The Orc6 subunit,

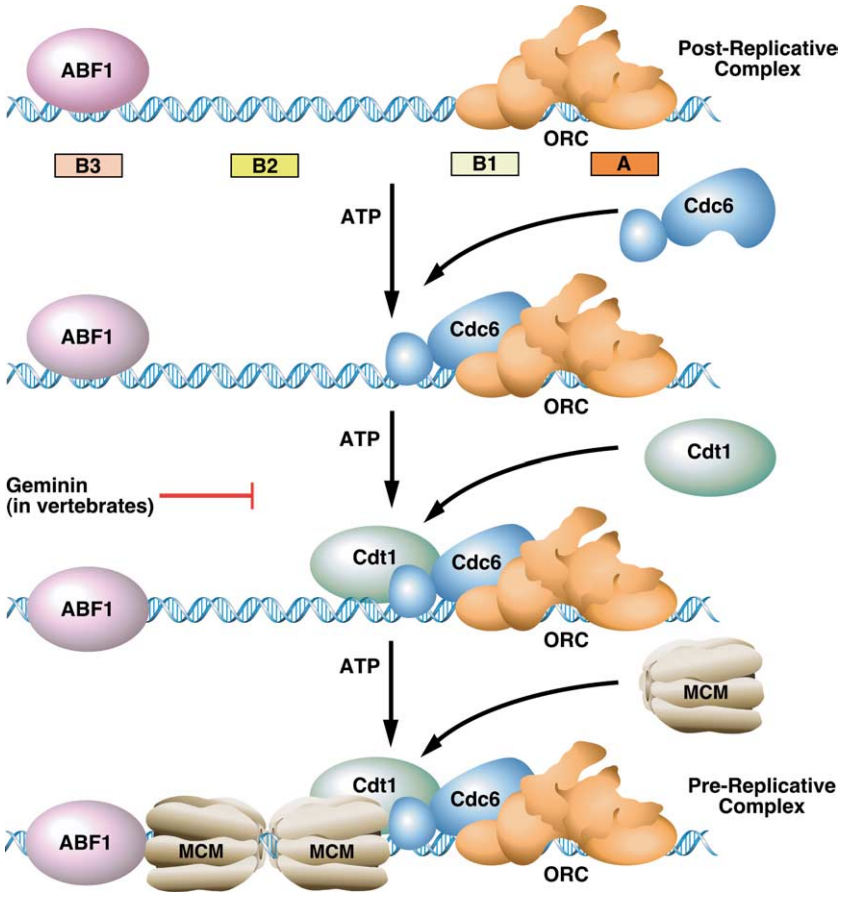

Fig. 2. Formation of the pre-Replication Complex in eukaryotes involves multiple AAA^+ proteins at the origins of DNA replication. ORC, Cdc6p and Cdt1p cooperate to load the MCM proteins that most likely function as a DNA helicase like SV40 T Ag. MCM structure is modeled from [38,100].

while localized to nuclear chromatin in interphase cells, binds to kinetochores in mitosis and also functions in a reticular network to promote cytokinesis [75,76]. Interestingly, Orc2 is also associated with centrosome throughout the cell division cycle and depletion of Orc2 from cells causes significant centrosomes defects [74]. Detailed studies have not been published for the other subunits, but it is likely that they also will participate in other parts of the chromosome replication and segregation cycle.

Another gene product that participates in origin recognition is the Cdc6 protein (Fig. 2). CDC6 was first identified as a mutant in the set of cell division cycle mutants described by Hartwell and colleagues and analysis of the phenotype of *CDC6* mutants showed that it had an execution point in late G1 just prior to entry into S phase [77]. The *S. pombe* Cdc6p homologue (Cdc18) was isolated as a regulator of DNA replication because overexpression of the protein caused cells to re-replicate in a single cell division cycle [78]. A link to origins of DNA replication was implied when overexpression of Cdc6p was identified as a suppressor of an *orc5-1* mutant [7]. Cdc6p was shown to control the frequency of initiation of DNA replication and certain mutants in *CDC6* cause an over-replication phenotype [79]. Interestingly, combining cyclin-CDK phosphorylation site mutants in both ORC subunits and Cdc6p with constitutive expression of MCM in the nucleus of cells also caused over-replication of the genome in a single cell cycle [80]. This is in part due to cyclin-CDK complex (Clb2p-Cdc28p) binding directly to Cdc6p in a phosphorylation-dependent manner and inhibiting Cdc6p activity [81].

Cdc6p binds directly to ORC [82], potentially via an interaction with Orc1p [83], and enhances the DNA binding specificity of ORC to origin sequences [82]. Furthermore, Cdc6p interaction with ORC promotes significant structural changes in ORC, with three subunits becoming very sensitive to protease digestion in the presence of Cdc6p and ATP, but not ADP [82]. It is possible that Cdc6p, a AAA^+ protein, may bind to the another AAA^+ protein subunit in ORC (such as Orc1p) and form a structure where adjacent AAA^+ subunits interact like the AAA^+ subunits in the DNA polymerase clamp loader protein complexes that exist in bacteria or eukaryotic cells [34,35]. For example, the eukaryotic RFC clamp loader protein requires ATP for loading the DNA polymerase clamp PCNA onto primer-template DNAs. The ATPase activity of one subunit of RFC is activated by an arginine finger residue in an adjacent AAA^+ subunit of RFC [34,35]. The potential similarity between Cdc6p and the clamp loaders has been discussed previously, since Cdc6p is required form loading the MCM proteins onto chromatin in vivo and in vitro [6,84–86]. In vitro, MCM loading requires ATP and both Cdc6p and ORC bound to DNA. Although Cdc6p may be the functional equivalent of the RFC loader at origins of DNA replication, it is more likely that the combined ORC-Cdc6p

protein complex [82] functions as the MCM loading complex, utilizing ATP just like RFC does to load PCNA [20,35].

Cdc6p is a AAA^+ protein that is highly similar in primarily amino acid sequence to the Orc1p subunit. Like Orc1p (and other ORC subunits Orc4p and Orc5p), Cdc6p has Walker A and B motifs that bind ATP and mutations in the Walker A motif block formation of the pre-RC and assembly of MCM proteins onto the DNA during mitotic exit and in G1 phase [6,83,85]. It is equally likely that Cdc6p will exert its ATPase activity when complexed with ORC. Since ORC has at least three AAA^+ proteins, their association with Cdc6p might activate Cdc6p ATPase activity in much the same way as other multi-subunit AAA^+ ATPase proteins do.

The role of Cdc6p in licensing DNA replication is conserved [5,16,18,87–90] and in the *Xenopus* early embryo is a rate limiting component for licensing of chromosomes for initiation of DNA replication [91]. Interestingly, in *Xenopus* oocytes, where origins are not determined by specific DNA sequences, Cdc6p stimulates ORC binding to chromatin [92]. Thus, there appears conservation of the ORC-Cdc6p interaction even though the details of how the two proteins interact to determine origin specificity in animal cell chromosomes are still not understood.

ORC and Cdc6p cooperate to load the MCM proteins onto DNA in and around the origins of DNA replication (Fig. 2). Multiple MCM complexes are believed to be loaded onto the DNA by a single ORC-Cdc6p complex [39]. The loading requires another protein called Cdt1p. This protein was originally identified in *S. pombe* as a regulator of the G1 to S phase transition and later shown to be required for MCM protein loading in both *S. pombe* and *Xenopus* egg extracts [10,11]. Cdt1p binds directly to the MCM protein complex and to Cdc6p, suggesting that it may act as a chaperone for bringing the MCM proteins to the origin. In vertebrates, Cdt1p is inhibited by a protein called geminin that is degraded by the anaphase promoting complex (APC) as cells exit from the metaphase stage of mitosis [93–96].

Once the MCM proteins are loaded onto DNA, the origins of DNA replication are competent for initiation of DNA replication, but replication still requires the cells to commit to a new round of cell division. Commitment to cell division promotes the activation of at least two protein kinase complexes, the S phase cyclin-CDKs and Cdc7-Dbf4 Both bind directly to and phosphorylate proteins in the pre-RC [2]. The role of Cdc7-Dbf4 is primarily to activate the MCM, since this kinase phosphorylates some MCM protein subunits and importantly, a mutation in the *MCM5* gene partially suppresses a complete deletion of the *CDC7* gene [92].

5. Origin recognition in Archaea

Archaea represent a third domain of living organisms that, although prokaryotes because they lack a nucleus, are quite distinct from bacteria. Determination of the sequence of many arachaeon genomes has demonstrated that the DNA replication machinery more closely resembles the proteins present in eukaryotes compared to bacteria [55,97]. Archaea have many of the proteins that were identified as being required for SV40 DNA replication, such as the single stranded DNA binding protein RPA, the DNA polymerase clamp PCNA and the ATP-dependent clamp loader RFC. Archaeon species also have some of the pre-RC components, including a MCM like DNA helicase that forms a double-hexameric structure remarkably like SV40 T Ag (Fig. 3) [37,38,98–100]. Although archaeon species do not have an ORC like that present in eukaryotes, they have one or more copies of proteins that are highly sequence related to Orc1p and Cdc6p, often called the Orc1p/Cdc6p protein [101–103]. Some species have only one such protein, but others have two or three and the recent sequence of a halobacterium species predicts nine Orc1p/Cdc6p related proteins [104]. It should be noted, however, that some species appear to lack an Orc1p/Cdc6p orthologue, suggesting that they may have an alternative mode of initiation of DNA replication, perhaps analogous to the bacterial RNase directed initiation [105].

The role of the Orc1p/Cdc6p protein appears to be both in origin recognition and MCM loading. As far as has been determined, many chromosomes and plasmids in archaea have a single origin or DNA replication, but there are clear examples where multiple origins have been predicted or indeed shown to exist [106–108]. Interestingly, these origin sequences physically map to sites adjacent to the gene(s) encoding the Orc1p/Cdc6p origin recognition proteins.

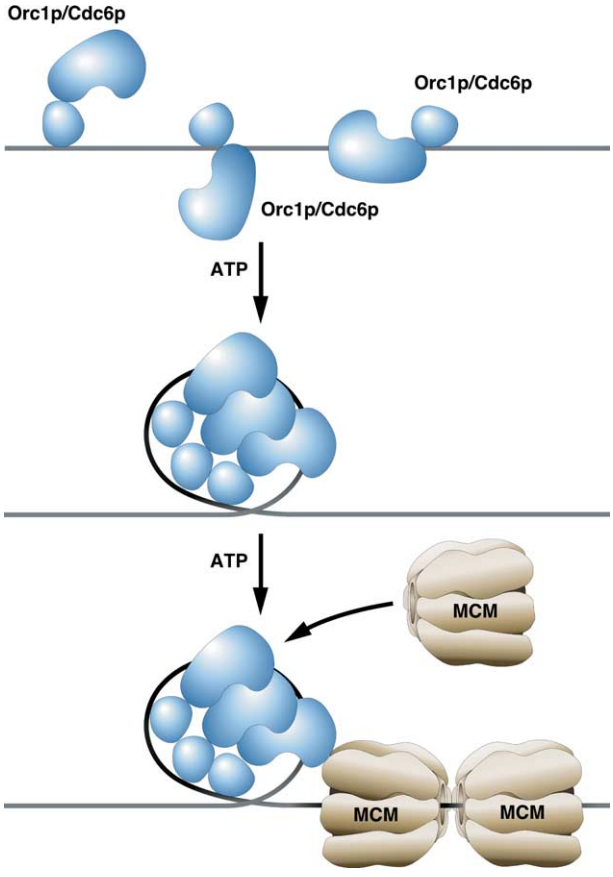

Fig. 3. Model for assembly of MCM proteins at archaeal origins of DNA replication. In this case, Orc1p/Cdc6p orthologues bind to specific sequences via a winged helix domain. Although not known, it is proposed that these proteins oligomerize to form an ORC like structure, or a structure similar to the multiple DnaA proteins that form at *E. coli oriC* in an ATP-dependent manner. The loading of the MCM helicase may not require a Cdt1p like protein and the multiple subunits of ORC that are present in eukaryotes. Cdc6 and DnaA structures are modeled from [109–111].

The three dimensional structure of Orc1p/Cdc6p from two archaeon species has shown it to contain a AAA^+ domain and a winged helix (WH) domain, both of which are predicted to be conserved in the eukaryotic Orc1p subunit of ORC and in Cdc6/Cdc18 (Fig. 3) [109,110]. The WH domain interacts with DNA and inhibits the DNA helicase activity of the MCM hexamer [89]. Most interestingly, the overall architecture of the archaeon Orc1p/Cdc6p proteins show striking resemblance to the structure of the bacterial DnaA initiator protein that also contains a AAA^+ domain and a helix-turn-helix domain that structurally resembles the WH domain from Orc1p/Cdc6p [111].

Recent studies have shown that the archaeon Orc1p/Cdc6p protein binds to specific DNA sequences at origins of DNA replication both in vivo and in vitro [89,103,112,113]. The WH domain is required for DNA binding, but unlike the eukaryotic ORC, ATP is not required. In one species, *Sulfoglobus solfataricus*, three Orc1p/Cdc6p orthologues exist and they interact with different sequences at two known and separate origins of DNA replication [106]. Two of the proteins, called Cdc6-1 and Cdc6-2, bind to the *oriC1* origin, whereas all three (Cdc6-1, Cdc6-2 and Cdc6-3) bind to the *oriC2* origin. These results suggest that the multiple Cdc6p proteins may oligomerize, perhaps in an ATP-dependent manner, at the origin and form a structure that resembles the multimers of DnaA that form in the presence of ATP at the *E. coli oriC* (Fig. 3). Thus, it appears that for origin recognition archaea may be a hybrid between the bacterial and eukaryotic modes of DNA replication initiation, although more biochemical studies are needed to investigate how the Orc1p/Cdc6p protein functions to recognize the origins of DNA replication and then load the MCM helicase.

The bacterial DnaA and archaeon Orc1p/Cdc6p proteins belong to a related clade of AAA^+ proteins that are distantly related to the RFC-like clamp loaders proteins, although there is little amino acid sequence similarity between the bacterial and archaeal proteins [114]. This group also includes the eukaryotic Orc1p, Orc4p and Orc5p proteins. Although this classification is based on structural organization of the proteins, all of them are involved in loading other proteins onto DNA and thus the structural relationship may reflect an evolutionarily related function in utilizing ATP to perform work to modify the structure of the target protein. For example, the RFC clamp loaders open the PCNA clamp to load it onto a primer-template DNA [20]. ORC and Cdc6 in eukaryotes load the MCM proteins and although the MCM proteins have a circular, hexameric structure, it is not yet clear whether they are loaded as a preformed complex or as individual subunits that form the hexamer on DNA. Whichever mechanism exists, it is the ATP-driven work that allows the correct assembly to be loaded onto DNA so that it can participate in the initiation of DNA replication.

6. ORC and beyond

Bacteria and archaea have single proteins that recognize the origins of DNA replication and assemble the DNA helicase onto the DNA prior to the initiation of DNA replication. This raises the question of why ORC in eukaryotes consists of six essential proteins that appear to perform the same function, and they even require other proteins such as Cdc6p and Cdt1p to load the MCM helicase. One possible explanation is that the temporal regulation of the initiation of DNA replication throughout S phase requires more coordination with cell cycle progression than in bacterial and archaeal cells. Furthermore, DNA replication in multi-cellular eukaryotes needs to be coupled with maintenance of genomic stability by accurately replicating every region of the genome only once per cell cycle. A single chromosome that fails to be replicated correctly will cause lethality and in an organism that accumulates altered copy numbers of genes in a single cell can result in a tumor and be a detriment to the organism as a whole.

In bacteria, chromosome segregation occurs at the same time as the DNA is being copied and there is not a requirement to condense the genome prior to separation of the daughter chromosomes. In contrast, eukaryotes have a different chromosome cycle because after chromosome duplication, the sister chromatids remain bound together prior to their coordinated segregation in mitosis. The segregation requires structures like centromeres and centrosomes, each of which is known to bind a subunit of the ORC [74–76]. In addition, formation of the pre-RC is essential for loading onto chromatin the SMC proteins that participate in sister chromatid cohesion and chromosome condensation [115,116]. ORC is also involved in higher order chromosome structure, since it associates with heterochromatin in both mammalian cells and *Drosophila* [117–120]. Thus, it is possible that the reason why ORC has six subunits is that they help coordinate chromosome duplication with aspects of chromosome structure and segregation. The complex structure of eukaryotic chromosomes may require an ATP-dependent complex that loads other proteins onto the DNA during each stage of the chromosome duplication and segregation cycle. ORC could fulfill this role.

Acknowledgments: Research in the author's laboratory is supported by grants from the US National Cancer Institute (CA13106) and the National Institute of General Medical Sciences (GM45436).

References

[1] Rao, P.N. and Johnson, R.T. (1970) Nature 225, 159–164.
[2] Bell, S.P. and Dutta, A. (2002) Annu. Rev. Biochem. 71, 333–374.
[3] Diffley, J.F. and Labib, K. (2002) J. Cell Sci. 115, 869–872.
[4] Bell, S.P. and Stillman, B. (1992) Nature 357, 128–134.
[5] Coleman, T.R., Carpenter, P.B. and Dunphy, W.G. (1996) Cell 87, 53–63.
[6] Donovan, S., Harwood, J., Drury, L.S. and Diffley, J.F. (1997) Proc. Natl. Acad. Sci. USA 94, 5611–5616.
[7] Liang, C, Weinreich, M. and Stillman, B. (1995) Cell 81, 667–676.
[8] Piatti, S., Bohm, T., Cocker, J.H., Diffley, J.F. and Nasmyth, K. (1996) Genes Dev. 10, 1516–1531.
[9] Santocanale, C. and Diffley, J.F. (1996) Embo J. 15, 6671–6679.
[10] Maiorano, D., Moreau, J. and Mechali, M. (2000) Nature 404, 622–625.
[11] Nishitani, H., Lygerou, Z., Nishimoto, T. and Nurse, P. (2000) Nature 404, 625–628.
[12] Tye, B.K. and Sawyer, S. (2000) J. Biol. Chem. 275, 34833–34836.
[13] Chong, J.P., Mahbubani, H.M., Khoo, C.Y. and Blow, J.J. (1995) Nature 375, 418–421.
[14] Labib, K., Tercero, J.A. and Diffley, J.F. (2000) Science 288, 1643–1647.
[15] Labib, K., Kearsey, S.E. and Diffley, J.F. (2001) Mol. Biol. Cell 12, 3658–3667.
[16] Gillespie, P.J., Li, A. and Blow, J.J. (2001) BMC Biochem. 2, 15.

[17] Aparicio, O.M., Weinstein, D.M. and Bell, S.P. (1997) Cell 91, 59–69.
[18] Madine, M.A., Swietlik, M., Pelizon, C., Romanowski, P., Mills, A.D. and Laskey, R.A. (2000) J. Struct. Biol. 129, 198–210.
[19] Nishitani, H. and Lygerou, Z. (2002) Genes Cells 7, 523–534.
[20] Waga, S. and Stillman, B. (1998) Annu. Rev. Biochem. 67, 721–751.
[21] Borowiec, J.A., Dean, F.B., Bullock, P.A. and Hurwitz, J. (1990) Cell 60, 181–184.
[22] Prives, C. (1990) Cell 61, 735–738.
[23] Wessel, R., Schweizer, J. and Stahl, H. (1992) J. Virol. 66, 804–815.
[24] Stenlund, A. (2003) Nat. Rev. Mol. Cell Biol. 4, 777–785.
[25] Fouts, E.T., Yu, X., Egelman, E.H. and Botchan, M.R. (1999) J. Biol. Chem. 274, 4447–4458.
[26] Chen, G. and Stenlund, A. (2002) Mol. Cell Biol. 22, 7712–7720.
[27] Gillitzer, E., Chen, G. and Stenlund, A. (2000) Embo J. 19, 3069–3079.
[28] Voitenleitner, C. and Botchan, M. (2002) J. Virol. 76, 3440–3451.
[29] Abbate, E.A., Berger, J.M. and Botchan, M.R. (2004) Genes Dev. 18, 1981–1996.
[30] Gomez-Lorenzo, M.G., Valle, M., Frank, J., Gruss, C., Sorzano, C.O., Chen, X.S., Donate, L.E. and Carazo, J.M. (2003) Embo J. 22, 6205–6213.
[31] Gai, D., Li, D., Finkielstein, C.V., Ott, R.D., Taneja, P., Fanning, E. and Chen, X.S. (2004) J. Biol. Chem. 279, 38952–38959.
[32] Gai, D., Zhao, R., Li, D., Finkielstein, C.V. and Chen, X.S. (2004) Cell 119, 47–60.
[33] Bradshaw, E.M., Sanford, D.G., Luo, X., Sudmeier, J.L., Gurard-Levin, Z.A., Bullock, P.A. and Bachovchin, W.W. (2004) Biochemistry 43, 6928–6936.
[34] Davey, M.J., Jeruzalmi, D., Kuriyan, J. and O'Donnell, M. (2002) Nat. Rev. Mol. Cell Biol. 3, 826–835.
[35] Bowman, G.D., O'Donnell, M. and Kuriyan, J. (2004) Nature 429, 724–730.
[36] Adachi, Y., Usukura, J. and Yanagida, M. (1997) Genes Cells 2, 467–479.
[37] Chong, J.P., Hayashi, M.K., Simon, M.N., Xu, R.M. and Stillman, B. (2000) Proc. Natl. Acad. Sci. USA 97, 1530–1535.
[38] Fletcher, R.J., Bishop, B.E., Leon, R.P., Sclafani, R.A., Ogata, C.M. and Chen, X.S. (2003) Nat. Struct. Biol. 10, 160–167.
[39] Edwards, M.C., Tutter, A.V., Cvetic, C., Gilbert, C.H., Prokhorova, T.A. and Walter, J.C. (2002) J. Biol. Chem. 277, 33049–33057.
[40] Laskey, R.A. and Madine, M.A. (2003) EMBO Rep. 4, 26–30.
[41] Blow, J.J. and Laskey, R.A. (1986) Cell 47, 577–587.
[42] Newlon, C.S. and Theis, J.F. (1993) Curr. Opin. Genet. Dev. 3, 752–758.
[43] Brewer, B.J. and Fangman, W.L. (1987) Cell 51, 463–471.
[44] Marahrens, Y. and Stillman, B. (1992) Science 255, 817–823.
[45] Chuang, R.Y., Chretien, L., Dai, J. and Kelly, T.J. (2002) J. Biol. Chem. 277, 16920–16927.
[46] Kong, D. and DePamphilis, M.L. (2002) Embo J. 21, 5567–5576.
[47] Lee, J.K., Moon, K.Y., Jiang, Y. and Hurwitz, J. (2001) Proc. Natl. Acad. Sci. USA 98, 13589–13594.
[48] Segurado, M., de Luis, A. and Antequera, F. (2003) EMBO Rep. 4, 1048–1053.
[49] Chuang, R.Y. and Kelly, T.J. (1999) Proc. Natl. Acad. Sci. USA 96, 2656–2661.
[50] Diffley, J.F. and Stillman, B. (1988) Proc. Natl. Acad. Sci. USA 85, 2120–2124.
[51] Beall, E.L., Manak, J.R., Zhou, S., Bell, M., Lipsick, J.S. and Botchan, M.R. (2002) Nature 420, 833–837.
[52] Thoma, F., Bergman, L.W. and Simpson, R.T. (1984) J. Mol. Biol. 177, 715–733.
[53] Diffley, J.F. and Cocker, J.H. (1992) Nature 357, 169–172.
[54] Diffley, J.F., Cocker, J.H., Dowell, S.J. and Rowley, A. (1994) Cell 78, 303–316.
[55] Bernander, R. (2003) Mol. Microbiol. 48, 599–604.
[56] Lee, D.G. and Bell, S.P. (1997) Mol. Cell Biol. 17, 7159–7168.
[57] Bell, S.P., Kobayashi, R. and Stillman, B. (1993) Science 262, 1844–1849.
[58] Lee, D.G., Makhov, A.M., Klemm, R.D., Griffith, J.D. and Bell, S.P. (2000) Embo J. 19, 4774–4782.
[59] Tugal, T., Zou-Yang, X.H., Gavin, K., Pappin, D., Canas, B., Kobayashi, R., Hunt, T. and Stillman, B. (1998) J. Biol. Chem. 273, 32421–32429.
[60] Klemm, R.D. and Bell, S.P. (2001) Proc. Natl. Acad. Sci. USA 98, 8361–8367.
[61] Takenaka, H., Makise, M., Kuwae, W., Takahashi, N., Tsuchiya, T. and Mizushima, T. (2004) J. Mol. Biol. 340, 29–37.
[62] Leatherwood, J., Lopez-Girona, A. and Russell, P. (1996) Nature 379, 360–363.
[63] Vas, A., Mok, W. and Leatherwood, J. (2001) Mol. Cell Biol. 21, 5767–5777.
[64] Weinreich, M., Liang, C., Chen, H.H. and Stillman, B. (2001) Proc. Natl. Acad. Sci. USA 98, 11211–11217.
[65] Wilmes, G.M., Archambault, V., Austin, R.J., Jacobson, M.D., Bell, S.P. and Cross, F.R. (2004) Genes Dev. 18, 981–991.
[66] Findeisen, M., El-Denary, M., Kapitza, T., Graf, R. and Strausfeld, U. (1999) Eur. J. Biochem. 264, 415–426.
[67] Li, C.J., Vassilev, A. and DePamphilis, M.L. (2004) Mol. Cell Biol. 24, 5875–5886.
[68] DePamphilis, M.L. (2003) Gene 310, 1–15.
[69] Li, C.J. and DePamphilis, M.L. (2002) Mol. Cell Biol. 22, 105–116.
[70] Mendez, J., Zou-Yang, X.H., Kim, S.Y., Hidaka, M., Tansey, W.P. and Stillman, B. (2002) Mol. Cell 9, 481–491.
[71] Natale, D.A., Li, C.J., Sun, W.H. and DePamphilis, M.L. (2000) Embo J. 19, 2728–2738.
[72] Ohta, S., Tatsumi, Y., Fujita, M., Tsurimoto, T. and Obuse, C. (2003) J. Biol. Chem. 278, 41535–41540.
[73] Tatsumi, Y., Ohta, S., Kimura, H., Tsurimoto, T. and Obuse, C. (2003) J. Biol. Chem. 278, 41528–41534.
[74] Prasanth, S.G., Prasanth, K.V., Siddiqui, K., Spector, D.L. and Stillman, B. (2004) Embo J. 23, 2651–2663.
[75] Prasanth, S.G., Prasanth, K.V. and Stillman, B. (2002) Science 297, 1026–1031.
[76] Chesnokov, I.N., Chesnokova, O.N. and Botchan, M. (2003) Proc. Natl. Acad. Sci. USA 100, 9150–9155.
[77] Hereford, L.M. and Hartwell, L.H. (1974) J. Mol. Biol. 84, 445–461.
[78] Kelly, T.J., Martin, G.S., Forsburg, S.L., Stephen, R.J., Russo, A. and Nurse, P. (1993) Cell 74, 371–382.
[79] Liang, C. and Stillman, B. (1997) Genes Dev. 11, 3375–3386.
[80] Nguyen, V.Q., Co, C. and Li, J.J. (2001) Nature 411, 1068–1073.
[81] Mimura, S., Seki, T., Tanaka, S. and Diffley, J.F. (2004) Nature 431, 1118–1123.
[82] Mizushima, T., Takahashi, N. and Stillman, B. (2000) Genes Dev. 14, 1631–1641.
[83] Wang, B., Feng, L., Hu, Y., Huang, S.H., Reynolds, C.P., Wu, L. and Jong, A.Y. (1999) J. Biol. Chem. 274, 8291–8298.
[84] Perkins, G. and Diffley, J.F. (1998) Mol. Cell 2, 23–32.
[85] Weinreich, M., Liang, C. and Stillman, B. (1999) Proc. Natl. Acad. Sci. USA 96, 441–446.
[86] Seki, T. and Diffley, J.F. (2000) Proc. Natl. Acad. Sci. USA 97, 14115–14120.
[87] Cocker, J.H., Piatti, S., Santocanale, C., Nasmyth, K. and Diffley, J.F. (1996) Nature 379, 180–182.
[88] Jang, S.W., Elsasser, S., Campbell, J.L. and Kim, J. (2001) Biochem. J. 354, 655–661.
[89] Matsunaga, F., Forterre, P., Ishino, Y. and Myllykallio, H. (2001) Proc. Natl. Acad. Sci. USA 98, 11152–11157.
[90] Stoeber, K., Mills, A.D., Kubota, Y., Krude, T., Romanowski, P., Marheineke, K., Laskey, R.A. and Williams, G.H. (1998) Embo J. 17, 7219–7229.
[91] Whitmire, E., Khan, B. and Coue, M. (2002) Nature 419, 722–725.
[92] Harvey, K.J. and Newport, J. (2003) J. Biol. Chem. 278, 48524–48528.
[93] Hodgson, B., Li, A., Tada, S. and Blow, J.J. (2002) Curr. Biol. 12, 678–683.
[94] Wohlschlegel, J.A., Dwyer, B.T., Dhar, S.K., Cvetic, C., Walter, J.C. and Dutta, A. (2000) Science 290, 2309–2312.
[95] Tada, S., Li, A., Maiorano, D., Mechali, M. and Blow, J.J. (2001) Nat. Cell Biol. 3, 107–113.
[96] Ballabeni, A., Melixetian, M., Zamponi, R., Masiero, L., Marinoni, F. and Helin, K. (2004) Embo J. 23, 3122–3132.

[97] Kelman, L.M. and Kelman, Z. (2004) Trends Microbiol. 12, 399–401.
[98] Kelman, Z., Lee, J.K. and Hurwitz, J. (1999) Proc. Natl. Acad. Sci. USA 96, 14783–14788.
[99] Shechter, D.F., Ying, C.Y. and Gautier, J. (2000) J. Biol. Chem. 275, 15049–15059.
[100] Pape, T., Meka, H., Chen, S., Vicentini, G., van Heel, M. and Onesti, S. (2003) EMBO Rep. 4, 1079–1083.
[101] De Felice, M., Esposito, L., Pucci, B., Carpentieri, F., De Falco, M., Rossi, M. and Pisani, F.M. (2003) J. Biol. Chem. 278, 46424–46431.
[102] De Felice, M., Esposito, L., Pucci, B., De Falco, M., Manco, G., Rossi, M. and Pisani, F.M. (2004) Biochem. J. 381, 645–653.
[103] De Felice, M., Esposito, L., Pucci, B., De Falco, M., Rossi, M. and Pisani, F.M. (2004) J. Biol. Chem. 279, 43008–43012.
[104] Ng, W.V., et al. (2000) Proc. Natl. Acad. Sci. USA 97, 12176–12181.
[105] Hendrickson, E.L., et al. (2004) J. Bacteriol. 186, 6956–6969.
[106] Robinson, N.P., Dionne, I., Lundgren, M., Marsh, V.L., Bernander, R. and Bell, S.D. (2004) Cell 116, 25–38.
[107] Zhang, R. and Zhang, C.T. (2003) Biochem. Biophys. Res. Commun. 302, 728–734.
[108] Contursi, P., Pisani, P.M., Grigoriev, A., Cannio, R., Bartolucci, S. and Rossi, M. (2004) Extremophiles 8, 385–391.
[109] Singleton, M.R., Morales, R., Grainge, I., Cook, N., Isupov, M.N. and Wigley, D.B. (2004) J. Mol. Biol. 343, 547–557.
[110] Liu, J., Smith, C.L., DeRyckere, D., DeAngelis, K., Martin, G.S. and Berger, J.M. (2000) Mol. Cell 6, 637–648.
[111] Erzberger, J.P., Pirruccello, M.M. and Berger, J.M. (2002) Embo J. 21, 4763–4773.
[112] Capaldi, S.A. and Berger, J.M. (2004) Nucleic Acids Res. 32, 4821–4832.
[113] Grainge, I., Scaife, S. and Wigley, D.B. (2003) Nucleic Acids Res. 31, 4888–4898.
[114] Iyer, L.M., Leipe, D.D., Koonin, E.V. and Aravind, L. (2004) J. Struct. Biol. 146, 11–31.
[115] Suter, B., Tong, A., Chang, M., Yu, L., Brown, G.W., Boone, C. and Rine, J. (2004) Genetics 167, 579–591.
[116] Takahashi, T.S., Yiu, P., Chou, M.F., Gygi, S. and Walter, J.C. (2004) Nat. Cell Biol. 6, 991–996.
[117] Loupart, M.L., Krause, S.A. and Heck, M.S. (2000) Curr. Biol. 10, 1547–1556.
[118] Pak, D.T., Pflumm, M., Chesnokov, L., Huang, D.W., Kellum, R., Marr, J., Romanowski, P. and Botchan, M.R. (1997) Cell 91, 311–323.
[119] Shareef, M.M., Badugu, R. and Kellum, R. (2003) Genetica 117, 127–134.
[120] Shareef, M.M., King, C., Damaj, M., Badagu, R., Huang, D.W. and Kellum, R. (2001) Mol. Biol. Cell 12, 1671–1685.